DOCUMENT RESUME

ED 189 159 TM 800 330

AUTHOR Reynolds, Cecil R.; Gutkin, Terry B.
TITLE Predictive Validity of the WISC-R for White and Mexican-American Children.
PUB DATE Apr 80
NOTE 17p.; Paper presented at the Annual Meeting of the American Educational Research Association (64th, Boston, MA, April 7-11, 1980).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Academic Achievement; Achievement Tests; *Anglo Americans; Emotional Disturbances; *Intelligence Tests; Intermediate Grades; *Mexican Americans; Multiple Regression Analysis; *Predictive Validity; Test Bias
IDENTIFIERS *Chicanos; *Wechsler Intelligence Scale for Children (Revised); Wide Range Achievement Test

ABSTRACT

Predictive validity of the Weschler Intelligence Scale for Children-Revised (WISC-R), was investigated with 174 Chicanos and 94 Anglos, ages 10-11, who were referred for psychological services. Wide Range Achievement Test Reading, Spelling, and Arithmetic subtest scores were each regressed on WISC-R Verbal, Performance and Full Scale Intelligence Quotients. Regression lines for the prediction of achievement were compared across race through the Potthoff analysis, which provides a simultaneous test of slope and intercept values. As expected, both groups scored substantially below norm on each measure. Subtest correlations were smaller than usual for both groups; they ranged from .03 to .56. Both sets of results were attributed to restriction of range. Of the nine Potthoff comparisons, only one differed significantly. It was concluded that the WISC-R demonstrated predictive validity across race. (Author/CP)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED189159

Predictive Validity of the WISC-R For White and Mexican-American Children

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Cecil R. Reynolds
Acting Director
Buros Institute of Mental Measurements
135 Bancroft Hall
Universtiy of Nebraska-Lincoln
Lincoln, Nebraska 68588

Terry B. Gutkin
Director
School Psychology Training Program
130 Bancroft Hall
University of Nebraska-Lincoln
Lincoln, Nebraska 68588

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
C. Reynolds

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

A paper presented to the annual meeting of the American Educational Research Association, Boston: April, 1980.

TM 800330

Abstract

Predictive validity of the Wechsler Intelligence Scale for Children-Revised was investigated across race for 174 Chicano and 94 Anglo children referred for psychological services. Regression lines for the prediction of achievement were compared across race through the Potthoff analysis, which provides a simultaneous test of slope and intercept values. Results of these comparisons generally supported the predictive validity of the WISC-R across race with this referral sample of young children.

The problem of bias in psychological and educational tests continues to be of great interest and importance to the general field of psychology. The use of psychological and educational tests with disadvantaged (or culturally different) students has come under close scrutiny in the last decade, primarily due to the now famous Larry P. case in San Francisco. Public Law 94-142 (The Education for All Handicapped Children Act of 1975) also mandates bias-free assessment of children for special education placement in the public schools. Even though most criticisms of testing have focused on the "misdiagnosis" of black children as mentally retarded (e.g., Mercer, 1973; Williams, 1970), empirical research in the area has focused on tests for personnel selection (e.g., Bartlett & O'Leary, 1969; Einhorn & Bass, 1971) and college admissions (e.g., Cleary, 1968, Kallingal, 1971). Results typically find the tests in question to be nonbiased in the prediction of minority performance.

The problem of bias is a complex one and no single definition has as yet been accepted (Flaugher, 1978; Hunter & Schmidt, 1976; Thorndike, 1971). In 1968, the American Psychological Association Board of Scientific Affairs appointed a committee to study the use of psychological and educational tests with the disadvantaged. In

reporting on this issue, the committee (Cleary, Humphreys, Kendrick, & Wesman, 1975) offered a definition of test bias. Content and construct validity were clearly considered to be important variables in the issue of bias by the committee. Nevertheless, the focus of the committee was on predictive validity:

> A test is considered fair for a particular use if the inference drawn from the test score is made with the smallest feasible random error and if there is no constant error in the inference as a function of membership in a particular group (Cleary et al., 1975, p. 25).

The definition offered by the APA Committee is a restatement of previous definitions by Cardall and Coffman (1964), Cleary, (1968), and Potthoff (1966).

Oakland and Matuszek (1977) examined special education class placement procedures under several proposed models of bias and demonstrated that the Cleary et al. (1975) regression model results in the smallest number of children being misplaced, though under certain legislative conditions they favored the Thorndike (1971) selection model. A variety of psychological and educational personnel have adopted the regression approach to test fairness including leading psychological authorities on testing (Anastasi, 1976; Cronbach, 1970; Humphreys, 1973), educational researchers (e.g., Kallingal, 1971), industrial/organizational psychologists (e.g, Bartlett & O'Leary, 1969), and even noted

critics of testing (e.g., Kirkpatrick, 1970). A statistical technique provided by Potthoff (1966) has also received widespread acceptance in the examination of regression lines to test bias under the above proferred definition (Schmidt & Hunter, 1974).

The Wechsler Intelligence Scale for Children-Revised (WISC-R; Wechsler, 1974) is presently the most widely employed intelligence test in the evaluation of children between the ages of 6 and 16½ years. Construct validity of the WISC-R now seems to be well established across race for normal (Gutkin & Reynolds, in press a Reschly, 1978; Vance & Wallbrown, 1978) and referred (Gutkin & Reynolds, in press b) children and also across sex (Reynolds & Gutkin, in press). Predictive validity of the WISC-R across race has only recently come under scrutiny. Reschly and Sabers (1979) examined the prediction of achievement by the WISC-R across race for normal Anglo, Black, Chicano, and Native American Papago children. For these four groups, slopes of the regression equations were generally equivalent though intercepts typically differed. When intercepts differed however, Anglo achievement scores were typically underpredicted, thus indicating bias against whites. Reynolds and Hartlage (1979) used the Potthoff (1966) technique to compare WISC and WISC-R Full Scale IQ regression equations, for the prediction of achievement, across race. These researchers demonstrated homogeneity of regression across race for the WISC-R and its predecess the 1949 WISC. Equivalent

predictive validity of the WISC-R across sex has also recently been demonstrated (Reynolds, Gutkin, Dappen, and Wright, 1979).

Although a picture of "fairness" is clearly developing with regard to construct and predictive validity of the WISC-R, the importance of the question of bias in psychological assessment demands more extensive investigation. The purpose of the present study was to investigate the predictive validity of the WISC-R across race for groups of abnormal White and Chicano children. The only previous study of bias in prediction with chicanos (Reschly & Sabers, 1979) involved children from regular classrooms with no known pathology of intellect, behavior, or personality.

Method

Subjects:

The children for the study consisted of the total referral population of Anglos and Chicanos from a school district in a moderately-sized urban area of the southwestern United States. From a total of 356 referrals to school psychological services by regular classroom teachers over the course of a full school year, complete data were available on 94 Anglos (M age = 11-1, SD = 2-9) and 174 Chicanos (M age = 10-6, SD = 2-6). While all socio-economic status groups were represented, the children were from primarily lower middle-class backgrounds.

Procedure

WISC-R Verbal, Performance, and Full Scale IQs were obtained

for each subject along with Wide Range Achievement Test (WRAT) data for achievement in Reading, Spelling, and Arithmetic. Each test was administered by a certified psychologist with no knowledge of the present study. The two tests were typically administered within several weeks of one another. Both are scaled to a mean of 100 and SD of 15 for their respective standardization sample. Since substantial restriction of range occurred within the Anglo and Chicano groups on the criterion measures, as is typical of referral populations, the scores (WISC-R and WRAT) for each racial group were normalized by converting them to z-scores with a mean of zero and SD of one. Pearson product-moment coefficients of correlation were then calculated between each dependent (WRAT) variable and each independent (WISC-R) variable for each group.

WRAT Reading, Spelling, and Arithmetic subtest scores were then each regressed on WISC-R Verbal IQs (VIQ), Performance IQs (PIQ), and Full Scale IQs (FSIQ). Regression lines for the prediction of achievement for each group were then contrasted by the Potthoff (1966) technique. The Potthoff technique is a sensitive analysis providing a single F-ratio that simultaneously tests regression coefficients (slopes) and intercept values. If a significant F results, slopes and intercepts may then be assessed separately to determine whether the resulting bias in prediction is constant (intercepts differ) or changes with the distance of scores from the mean (slopes differ).

Slopes and intercepts must both be equivalent prior to concluding homogeneity of regression across groups. Only when slopes and intercepts are the same can a common regression equation (derived by combining the groups in question) be applied. If homogeneity of regression across groups does not occur, then in order to have fair-use of test scores, separate equations for each group must be employed.

Results and Discussion

Means and SDs for the Anglo and Chicano groups on each measure are reported in table 1. As is routinely found, both groups

---

Insert Table 1 About Here

---

scored substantially below the normative sample on each measure. Restriction of range occurred most prominently on the three WRAT subtests. Correlations of each WISC-R IQ with each WRAT subtest are presented in Table 2. These correlations are smaller than are typically found for both groups, probably due to the restriction of

---

Insert Table 2 About Here

---

range previously noted. If consistently replicated with other referral or abnormal populations, the IQ may be of questionable utility in predicting the academic performance of these groups.

Of the 9 Potthoff comparisions conducted between the regression

lines of each racial group (e.g., WISC-R VIQ prediction of WRAT Reading for Anglo versus Chicano children), only the regression equations for the WISC-R PIQ and WRAT Arithmetic subtest differed significantly ($F = 3.22$, $p < .05$). Although the slope of the two regression lines was similar, the intercepts differed substantially, and in such a direction that application of a common regression equation results in overprediction of the Chicano scores and underprediction of Anglo scores. The finding of one significant difference in nine such sensitive comparisions can, however, reasonably be attributed to chance.

The present results lend support to the existance of a consistent relationship between IQ and academic performance independent of the child's race, at least within referral populations. Prior to concluding that the WISC-R, or any other measure, is bias-free (under the regression definition) more such studies will be needed using a variety of criterion measures. Validation against a variety of measures, including other achievement tests, teacher-made tests, and class grades, must all occur since we are always plagued by the question of bias in the criterion measure. More attention should also be applied to the sex variable in studies of differential validity, especially since most special education programs are at present much more blatantly sexist (males outnumber females 4:1) than racist in composition (Lambert, 1979).

References

Anastasi, A. Psychological testing, 4th edition. New York: Macmillan, 1976.

Bartlett, C. J., & O'Leary, B. S. A differential prediction model to moderate the effect of heterogeneous groups in personnel selection and classification. Personnel Psychology, 1969, 22, 1-18.

Cardall, C., & Coffman, W. A method for comparing the performance of different groups on the items in a test. Research and Development Reports, 1964, 64-5 no. 9, College Entrance Examination Board.

Cleary, T. A. Test bias: Prediction of grades of negro and white students in integrated colleges. Journal of Educational Measurement, 1968, 5, 115-124.

Cleary, T. A., Humphreys, L. G., Kendrick, S. A. & Wesman, A. Educational uses of tests with disadvantaged students. American Psychologists, 1975, 30, 15-41.

Cronbach, L. J. Essentials of psychological testing. New York: Harper & Rowe, 1970.

Einhorn, J. J. & Bass, A. R. Methodological considerations relevant to discrimination in employment testing. Psychological Bulletin, 1971, 75, 261-269.

Flaugher, R. L. The many definitions of test bias. American Psychologist, 1978, 33, 671-679.

Gutkin, T. B. & Reynolds, C. R. Factorial similarity of the WISC-R

for blacks and whites in the standardization sample. Journal of Educational Psychology, in press. a.

Gutkin, T. B., & Reynolds, C. R. Factorial similarity of the WISC-R for Anglos and Chicanos referred for psychological services. Journal of School Psychology, in press. b.

Humphreys, L. G. Statistical definitions of test validity for minority groups. Journal of Applied Psychology, 1973, 58, 1-4

Hunter, J. E. & Schmidt, F. L. Critical analysis of the statistical and ethical implications of various definitions of test bias. Psychological Bulletin, 1976, 83, 1053-1071.

Kallingal, A. The prediction of grades for black and white students at Michigan State University. Journal of Educational Measurement, 1971, 8, 263-265.

Kirkpatrick, J. J. The psychological testing establishment: Vested interests versus responsibility. Paper presented to the annual meeting of the American Psychological Association, Miami Beach: September, 1970.

Lambert, N. What of tomorrow's children: Tests on Trial II. Paper presented to the annual meeting of the National Association of School Psychologists, San Diego; March, 1979.

Mercer, J. R. Labelling the mentally retarded. Berkeley, California: University of California Press, 1973.

Oakland, T., & Matuszek, P. Using tests in non-discriminatory assessment. In T. Oakland (Ed.), Psychological and educational assessment of minority group children.

Potthoff, R. F. Statistical aspects of the problem of biases in psychological tests. Institute of Statistics Mimeo Series No. 479, Chapel Hill, N. C.: Department of Statistics, 1966.

Reschly, D. J. WISC-R factor structures among Anglos, Blacks, Chicanos, and Native American Papagos. Journal of Consulting and Clinical Psychology, 1978, 46, 417-422.

Reschly, D. J. & Sabers, D. Analysis of test bias in four groups with the regression definition. Journal of Educational Measurement, 1979, 16, 1-9.

Reynolds, C. R. & Gutkin, T. B. Stability of the WISC-R factor structure across sex at two age levels. Journal of Clinical Psychology, in press.

Reynolds, C. R. & Gutkin, T. B., Dappen, L., & Wright, D. Differential validity of the WISC-R for boys and girls referred for psychological services. Perceptual and Motor Skills, 1979, 48, 868 - 870.

Reynolds, C. R. & Hartlage, L. C. Comparison of WISC and WISC-R regression lines for academic predition with black and white referred children. Journal of Consulting and Clinical Psychology, 1979, in press.

Schmidt, F. L. & Hunter, J. E. Racial and ethnic bias in psychological tests: Divergent implications of two definitions of test bias. American Psychologist, 1974, 29, 1-8.

Thorndike, R. L. Concepts of culture-fairness. Journal of Educational Measurement, 1971, 8, 63-70.

Vance, H. B. & Wallbrown, F. H. The factor structure of intelligence for black children: A hierarchial approach. The Psychological Record, 1978, 28, 31-39.

Wechsler, D. W. Manual for the Wechsler Intelligence Scale for Children-Revised. New York: The Psychological Corporation, 1974.

Williams, R. L. Danger:Testing and dehumanizing black children. Clinical Child Psychology Newsletter, 1970, 9, 5-6.

Table 1

IQ and Achievement Scores for Anglo and Chicano Groups

	Full Scale IQ		Verbal IQ		Performance IQ		WRAT Reading		WRAT Spelling		WRAT Arithmetic	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Anglos N=94	80.91	15.01	79.14	16.25	85.52	14.97	80.34	12.45	78.16	9.67	77.28	10.60
Chicanos N=174	74.83	11.23	70.21	11.53	83.47	17.97	76.58	8.76	75.63	8.17	78.00	8.74

Table 2

WISC-R and Achievement Correlations for Anglos and Chicanos Referred for Psychological Services

WISC-R	WRAT,		
Anglos N=94	Reading	Spelling	Arithmetic
VIQ	.56	.44	.41
PIQ	.28	.15	.43
FSIQ	.48	.33	.47
Chicanos N=174			
VIQ	.42	.35	.41
PIQ	.03	.03	.25
FSIQ	.25	.18	.38